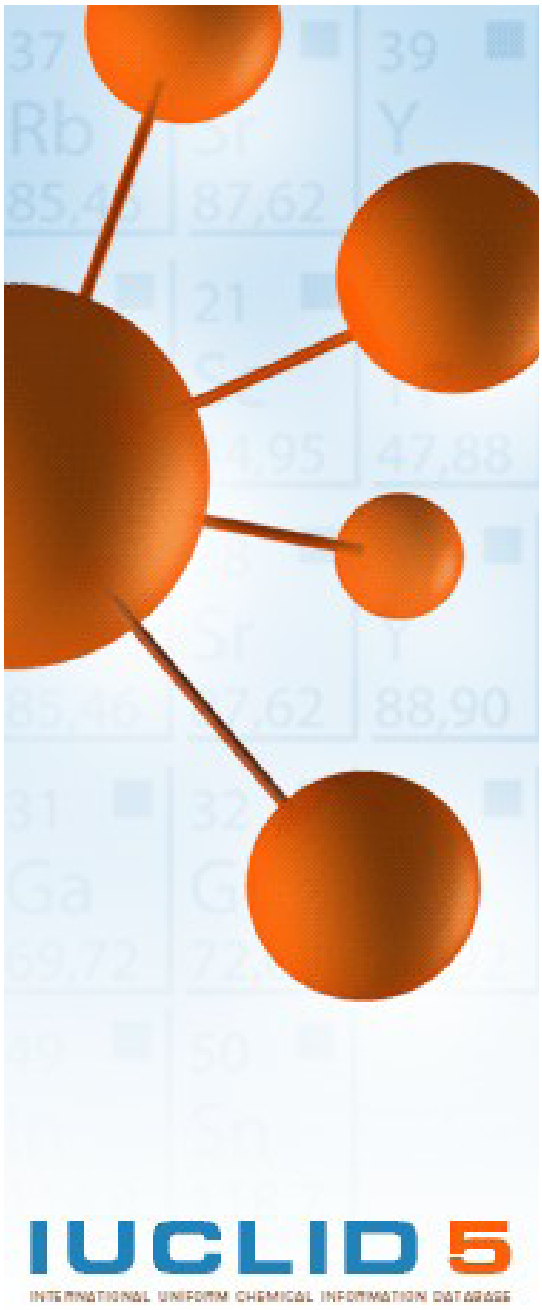


IUCLID 5
Guidance and Support

IUCLID 5
Query Tool Plugin
User Manual

January 2009
version 1.0

IUCLID 5 has been developed by the European Commission in association with the OECD

Table of contents

1 Overview....1
1.1 Query Indexer....1
1.2 Query Executor....2
2 Important note: requirements....2
3 Administration....3
3.1 First steps....4
3.2 Configuring the indexer....4
3.3 Creating the initial index....6
3.4 Disabling the index....7
3.5 Indexer status....8
3.6 Important notes....9
4 Executing Queries....9
4.1 Overview....10
4.2 Create a query....10
4.2.1 Defining query blocks....10
4.2.2 Combining query blocks....13
4.2.3 Defining the data layer....15
4.2.4 Defining the query result....15
4.3 Count a query....15
4.4 Execute a query....16
4.5 Save a query....17
4.6 Load a query....18
4.7 Delete a query....19
4.8 Export a query....19
4.9 Import a query....20
4.10 Important notes....20
5 Appendix....21
5.1 Units....21
5.2 Querying in numeric ranges....22
5.2.1 Ranges in IUCLID 5....22
5.2.2 Specifying the search criteria....23
5.2.3 The range search operators....25
5.3 Wildcards and escaping of characters....25

1 Overview

IUCLID 5 has a large and complex data model:

- thousands of data fields
- different documents linked by UUIDs
- different data types
- nested (cascading) data structures

The IUCLID 5 Query Tool delivers an extended query possibility that provides simple access to a defined set of business-relevant information.

The IUCLID 5 Query Tool consists of two types of plug-ins - a server-side and a client-side plug-in - that are co-operating to provide extended query functionality. Both plug-ins have to be installed to ensure proper operation of the IUCLID 5 Query Tool.

The IUCLID 5 Query Tool supports both, the distributed and the standalone version of IUCLID 5. However, the tool is optimized for the distributed version. There are some restrictions on installing and operating the IUCLID 5 Query Tool on the standalone version of IUCLID 5 (see "3.2 Configuring the indexer" for details).

The IUCLID 5 Query Tool uses a Query Indexer, which tracks all changes made to IUCLID 5 data, then extracts, and prepares the data for fast and convenient search operations. The Query Indexer creates an index table in the IUCLID 5 database. The Query Executor uses the index table to find IUCLID 5 data that matches the query criteria entered by the user.

1.1 Query Indexer

The Query Indexer consists of two basic parts, the query index and the query backlog. Whenever IUCLID 5 documents, e.g. substances, endpoints or mixtures are created, updated or deleted, the unique identifier (UUID) of these documents are stored in the query backlog, together with the action (delete, update, create) that has been executed on the data. The items contained in the query backlog are then processed and the processing result (fields/data items extracted from the documents) are then stored in the query index. This processing takes place in a configurable time interval. A time scheduler activates the indexer. The indexer then removes processed/indexed items from the backlog.

Users can find only indexed data through a query, i.e. it is desirable to use a short indexing interval and run the indexer frequently. However, indexing is a CPU time and memory consuming process, so the administrator has to find a good balance between a fast updating index and a good performance of the overall IUCLID 5 system (see "3.2 Configuring the indexer" for details).

The index only stores data in English language, so the users can only find data (especially phrases) in English.

All data is stored in upper case in the index tables. The indexer converts standard characters "a, b, c …, z" to upper case. Special language specific characters are not converted to upper case, e.g. á, â, ã, ä, å, æ, ç, è, é, ê, ë, ì, í, î, ï, ð, ñ, ò, ó, ô, õ, ö, ø, ù, ú, û, ü, ý, þ, ÿ. The exception is Smiles and InChI. This search pattern is case sensitive.

The Query Indexer indexes numeric values in standard units.

1.2 Query Executor

The Query Executor uses the query index to find the data that matches the query criteria entered by the user. The Query Executor finds only data that is already contained in the index. Depending on the indexer configuration (see "3.2 Configuring the indexer" for details), it can take up to one day before the data is available. On very powerful systems with less writing activities (import, edit, create dossier, delete) it is possible update the index tables nearly immediately. However, the minimal time interval for indexer runs is one minute. This minimizes the necessary update cycles.

The Query Executor converts search patterns into uppercase. The exception is Smiles and InChI. This search pattern is case sensitive.

The Query Executor converts numeric fields with units to standard units and queries the index with the standard units (see "5.1 Units").

2 Important note: requirements

Please be aware that the installation of the IUCLID 5 query tool requires that the twice disk space for the IUCLID 5 database is doubled. The IUCLID 5 FAQ recommends 5 MB per stored substance or dossier (without attachments). With an installation of the IUCLID 5 query tool, 10 MB per stored substance or dossier (without attachment) are required.

The query tool will need additional memory on an IUCLID 5 server or on an IUCLID 5 workstation. The users and administrators should be aware of the fact that the result of the queries executed by the query tool is in no way restricted, i.e. there might be ten thousands of items in a result set. The IUCLID 5 server or the IUCLID 5 workstation have to keep the biggest block result of a query (i.e. the biggest result of a single block or multiple blocks combined with "OR") in memory at some point in time (see "4.2 Create a query"). When the query has completed and the blocks are combined, the end-result has to be kept in memory. Please be aware, that when an out-of-memory situation in the an IUCLID 5 server or on an IUCLID 5 workstation occurs, it will not be isolated to the query tool, but also affect the rest of the IUCLID 5 system. This will also mean that users will loose unsaved data. If the out of memory situation occurs in the IUCLID 5 server, not only the user who caused the out of memory situation is affected, but also all other users of the system.

- IUCLID 5 workstation: A IUCLID 5 workstation has only one concurrent user, so the amount of additional memory needed only depends on the executed queries. The size of the single query blocks as well as the size of the end-result of a query determine the amount of needed memory. You should make at least 1GB of memory available to the IUCLID 5 workstation. This has to be done via configuration of the Java Virtual Machine. IUCLID 5 does not automatically use all available system memory. It uses only the memory made available via configuration.
- IUCLID 5 server: The number of users executing queries at the same time as well as the size of the single query blocks and the size of the end-result of a query de-

termine the amount of needed memory. You should make at least 1GB of memory available to the IUCLID 5 workstation. This has to be done via configuration of the Java Virtual Machine. IUCLID 5 does not automatically use all available system memory. It uses only the memory made available via configuration. As a rule of thumb, 100MB of memory should be made available per concurrent user (concurrently executing queries) of the query tool. This assumes that a query returns a maximum of 10000 Hits per end-result.

3 Administration

The "SuperUser" is the only user who can launch the IUCLID 5 Query Tool administration screen. Normal users, even when assigned to the administrator role, cannot launch the administration user interface. Indexing operations and their configuration can have a significant impact on the overall performance of IUCLID 5. Because of that, the administration screen has been restricted to the "SuperUser".

The administration interface launches from the IUCLID 5 task panel using the "Administration" link. It is included in the Query Tool group.

Task Panel: Query Tool Plug-Ins

The "SuperUser" uses the administration screen to

- configure the indexer
- create the initial index and start the indexer
- drop the index and stop the indexer
- browse status information on the indexer

3.1 First steps

The "SuperUser" has to configure the query indexer and then create the initial index, so that users can execute queries and get meaningful results from the IUCLID 5 Query Tool.

The "SuperUser" should execute the following steps to make a query index available for searching:

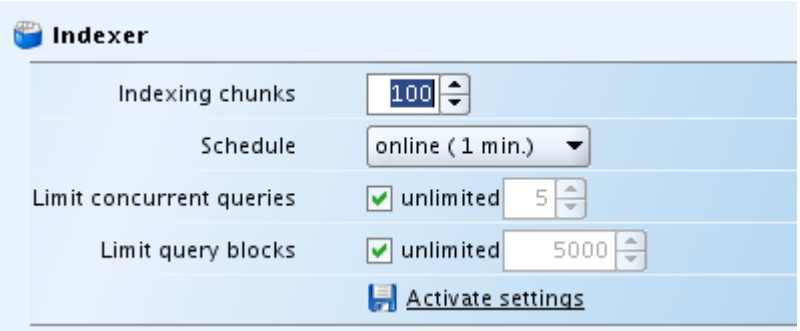
1. Configure the indexer (see "3.2 Configuring the indexer") to use indexing chunks of size 20 and set the schedule of the indexer to "Offline (at 23:15)".
2. After working hours, before a weekend, the "SuperUser" should create the initial index (see "3.3 Creating the initial index"). The "SuperUser" should do this when there is no load on the IUCLID 5 database (after working hours), because this task will put a heavy load on the IUCLID 5 system.
3. When the indexing mode changes to "normal" (see "3.3 Creating the initial index"), the "SuperUser" should switch the indexer configuration to use indexing chunks of size 20 and set the schedule of the indexer to "Online (5 min.)". The "SuperUser" should then monitor the statistic value "**Last indexing execution time**" (see "3.5 Indexer status") during working hours, i.e. when there is an average load on IUCLID 5.
    a. If the observed time is longer than 5 minutes, the "SuperUser" should adjust the indexer schedule, so that the timing interval is bigger than the time needed for indexing (the indexer will skip indexer runs if the timing interval is not big enough).
    b. If the observed time is shorter than 1 minute, the "SuperUser" should adjust the indexer schedule, so that the timing interval is "Online (1 min.)".
    c. If the observed time is between 1 minute and 5 minutes, the "SuperUser" should leave the indexer schedule at "Online (5 min.)".
4. The "SuperUser" might try different values for indexing chunks to fine tune the indexing (see "3.2 Configuring the indexer").

3.2 Configuring the indexer

The "SuperUser" configures the indexer in the "Indexer" section of the Query Tool administration screen.

The "SuperUser" can configure the size of the indexing chunks and the schedule of the indexer. Additionally, the "SuperUser" can restrict the number of concurrent queries that can be run by users and limit the size of the query blocks.

The limitation of the concurrent number of queries and the size of the query blocks lowers the memory consumption on the server. Note that it does not make sense to limit the concurrent number of queries on a IUCLID 5 workstation, because there will not be more than one concurrent query on a IUCLID 5 workstation. When the "SuperUser" limits the concurrent number of queries, users that run a query when this limit is reached, will receive an error message. When the "SuperUser" limits the size of the query blocks, users that run a query that has more result items than configured will receive an informational message that tells the users that query items were skipped.



Query Tool Administration: Configuring the indexer

During an indexer run, the indexer processes the complete backlog. After an indexer run, the backlog is empty. The indexer processes the backlog in chunks, to reduce the memory load on the IUCLID 5 system. When the time scheduler activates the indexer, the indexer takes the number of items defined in the indexing chunks setting from the backlog and processes these items. The indexer repeats this step until no backlog items remain in the backlog.

The indexing chunks setting is in the range from 20-100. A higher setting consumes more memory and processing power on the IUCLID 5 system (on the server or the workstation), but speeds up indexing. The "SuperUser" should only change the indexing chunks setting if optimisation is needed, otherwise the default setting should be used.

The "Schedule" setting defines the time interval in which the scheduler invokes the indexer. The possible settings are:

Online (1 min.)	**The scheduler invokes the indexer every minute.**
Online (5 min.)	The scheduler invokes the indexer every five minutes.
Online (10 min.)	The scheduler invokes the indexer every ten minutes.
Offline (at 20:15)	The scheduler invokes the indexer at 20:15 every day.
Offline (at 22:15)	The scheduler invokes the indexer at 22:15 every day.
Offline (at 23:15)	The scheduler invokes the indexer at 23:15 every day.
Offline (at 02:15)	The scheduler invokes the indexer at 02:15 every day.
Offline (at 04:15)	The scheduler invokes the indexer at 04:15 every day.
Offline (at 06:15)	The scheduler invokes the indexer at 06:15 every day.

Query Tool Indexer Schedule

Please be aware that if an indexer run takes so long that it is not completed until the scheduler tries to trigger the following indexer run, the scheduler skips the following run. Example: If the scheduler is configured for "Online (1 min.)" and indexing takes longer than one minute, the configured schedule cannot be reached. The "SuperUser" should avoid such a configuration and choose an interval that is big enough for the indexer runs to complete. Depending on the expected number of items in the backlog and the performance of the IUCLID 5 system, the "SuperUser" should choose a fast "online" index update or go for a nightly index update ("offline"). The values provided in the "Status" section can be used to analyse the system and find an appropriate indexer setting. The status field "Last indexing execution time" provides the time needed for the last indexer run.

Note that if the indexing schedule is too fast for the performance of the IUCLID 5 system, the whole IUCLID 5 system will slow down. The indexer settings can be changed at any time, even when the initial index was already created and the indexer is running.

Using the indexer on the distributed version

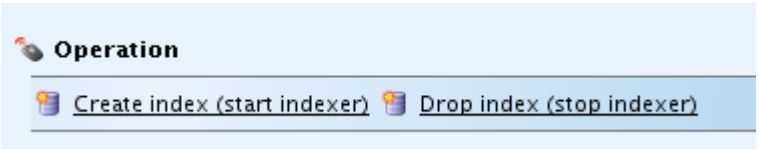
The indexer and the time scheduler that invokes the indexer runs on the IUCLID 5 server, i.e. as long as the IUCLID 5 server is running, indexing can take place. The indexer only consumes CPU time and memory on the server and puts no load on the client.

Using the indexer on the standalone version

It is not recommended to use the IUCLID 5 Query Tool on the standalone version that contains a huge amount of data (>200 substances/dossiers). Because the scheduler that invokes the indexer is part of IUCLID 5, it can only operate when the IUCLID 5 system is running. For a standalone version this means, that the indexer can only be invoked when the IUCLID 5 application is active. If for example the indexer is configured as "Offline (at 20:15), the user has to make sure that the IUCLID 5 application is running at that time. An additional drawback of operating the Query Tool in the standalone version is that when documents are updated or created they are placed in the backlog and subsequently indexed. On a standalone version, this slows down the IUCLID 5 application, whereas in the distributed version these are operations that are only happening on the server. The distributed client will not be affected by the time consuming indexing process.

3.3 Creating the initial index

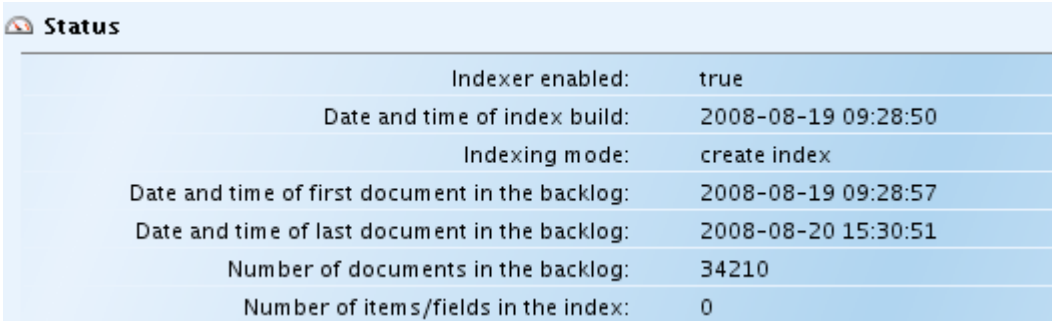
The "SuperUser" has to create the initial index and start the indexer after the installation of the IUCLID 5 Query Tool. The "SuperUser" creates the index in the "Operation" section of the Query Tool administration screen.



Query Tool Administration: Create index

When the "SuperUser" selects the "Create index (start indexer)" operation, the IUCLID 5 Query Tool starts to build (or rebuild) the complete index. First, the IUCLID 5 Query Tool scans the IUCLID 5 database and puts every document (marked for indexing) in the backlog. The indexer processes the content of the backlog during its normal indexer run as configured. Note that the indexing will only start after the process of filling the backlog has completed.

The "SuperUser" can monitor the index build in the "Operation" section of the Query Tool administration screen.


Status
Indexer enabled: true
Date and time of index build: 2008-08-19 09:28:50
Indexing mode: create index
Date and time of first document in the backlog: 2008-08-19 09:28:57
Date and time of last document in the backlog: 2008-08-20 15:30:51
Number of documents in the backlog: 34210
Number of items/fields in the index: 0

Query Tool Administration: Status

When the "Create index (start indexer)" operation is triggered, the "Indexing mode" changes (from "off") to "create index". The index is updated when the "Indexing mode" is displayed as "normal". The backlogging architecture of the indexer ensures that from the click on "Create index (start indexer)", every operation made on documents is tracked and the index will stay in sync with the IUCLID 5 database.

The index creation can run unattended. The initial creation of the index takes a lot of time. The initial index creation will take a few hours up to a few days (for IUCLID 5 databases that contain many documents). It is a good idea to run the initial index creation over a weekend, as it will slow down the IUCLID 5 system. Small IUCLID 5 databases will create the initial index in a few hours.

Do not stop the initial index creation before the first index item was created, e.g. by exiting the IUCLID 5 application on the standalone version, or by stopping the IUCLD 5 server on the distributed version. This will result in an invalid index. As soon as the first item in the index has been created ("Number of items/fields in the index" > 0), the IUCLID 5 indexer might be stopped, and will continue indexing when it is restarted. However, if the index creation was started and the indexer is interrupted before the creation of the first index item, i.e. the indexer is still in the process of filling the backlog, no valid index will be available!

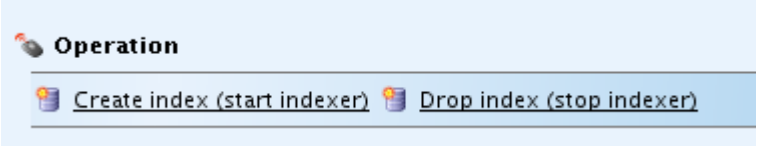
The index can only be dropped with "Drop index (stop indexer)", when the "Indexing mode" is "normal". Dropping the index is not possible until it was successfully created.

3.4 Disabling the index

After the indexer has created the intial index, the "SuperUser" can stop the indexer if necessary. If the "SuperUser" stops the indexer, it will be necessary to create the initial index again. The "SuperUser" should stop the indexer, if a restore operation of a IUCLID 5 database backup is done. This is necessary, as a database restore will deliver documents in random order without caring for the dependencies between

those documents. The indexer will not be able to keep the index in sync with the database because some of the dependencies between the documents will not be resolvable.

The "SuperUser" drops the index and stops indexing in the "Operation" section of the Query Tool administration screen.



Query Tool Administration: Create index

When the "SuperUser" selects the "Drop index (stop indexer)" operation, the IUCLID 5 Query Tool deletes the index table and stops the indexer. The IUCLID 5 Query Tool deletes the index table because it will go out-of-sync as soon the indexer stops.

The "SuperUser" can only stop the indexer after the indexer has finished the initial index creation. The "SuperUser" cannot use the "Drop index (stop indexer)" operation to abort the initial index creation!

3.5 Indexer status

The "Status" section of the Query Tool administration screen gives some status information about the indexer process and some statistical information on executed queries.

Status value	Description
Indexer enabled	Shows if the indexer is enabled and running.
Date and time of index build	The date and time when the index creation was triggered.
Indexing mode	Shows the operation mode of the indexer. Possible values are "off", "create index", "normal". "off": Indexer is not running "create index": Initial index creation in progress "normal": Initial index was created and indexer is now tracking document changes
Date and time of first document in the backlog	The date and time when the first (oldest) document in the backlog was added to the backlog
Date and time of last document in the backlog	The date and time when the last (newest) document in the backlog was added to the backlog.

Status value	Description
Number of documents in the backlog	The number of documents in the backlog that have to be processed during the next scheduled indexer run.
Date and time of first item/field in the index	The date and time of the oldest indexed item.
Date and time of last item/field in the index	The date and time of the newest indexed item.
Numbers if items/fields in the index	The size of the index
Last indexing execution time	The time the last indexer run needed.
Average query time	The average time a query took to execute (includes SQL queries and query block combination)
Average SQL query time	The average time SQL queries needed for execution. Together with "Average query time", one can see how much time is needed for combining blocks.
Average SQL query result size	The average size of results returned by SQL queries (this is the size of single query blocks).

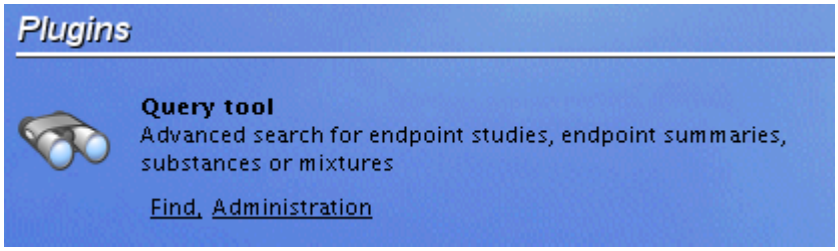
3.6 Important notes

Some important notes on indexer administration:

- Always disable the indexer before starting a restore operation in IUCLID 5
- Be aware that the initial indexing takes a lot of time and resources and cannot be interrupted. Only create the initial index if you are sure that the time is available and you will not do performance intensive operations during the index creation.

4 Executing Queries

Users build and execute queries through the IUCLID 5 Query Tool screen. The IUCLID 5 Query Tool screen launches from the IUCLID 5 task panel using the "Find" link. It is included in the Query Tool group.



Task Panel: Query Tool Plug-Ins

Other ways to launch the IUCLID 5 Query Tool screen is the toolbar icon or the "Plugins|Query Tool" menu.

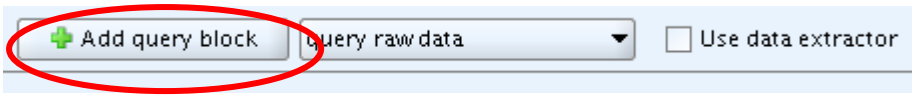
4.1 Overview

The user first has to create a query based on the defined query fields. If the query is valid, the user can execute the query. The queries can be stored and loaded or exported and imported.

4.2 Create a query

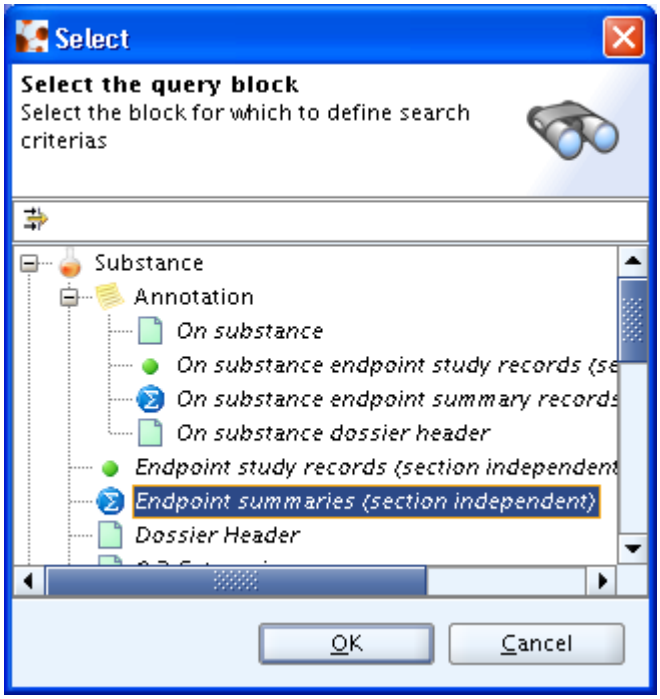
4.2.1 Defining query blocks

The user can define up to ten query blocks and combine these query blocks with the operations "AND", "OR" and "NOT". The Query Tool uses each of the query blocks to find field values in the query index. After searching the index, the results of the query blocks are combined.



Query Tool screen: Create a new query

The user first adds the necessary query blocks to create a query. The query blocks can be added by pressing "Add query block".

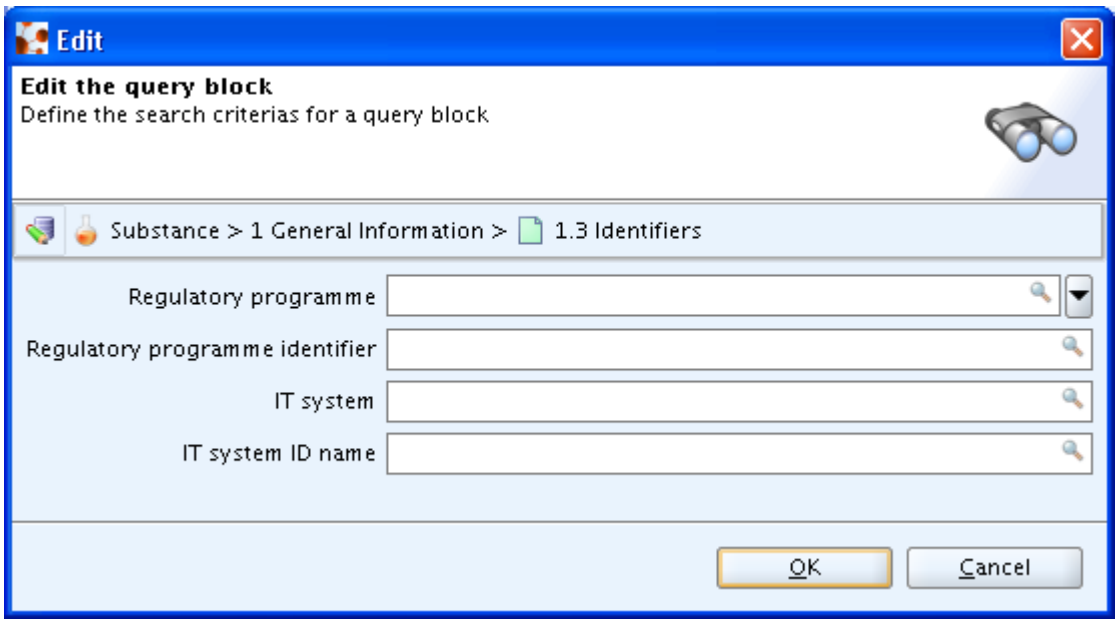


Query Tool screen: The queries

"Add query block" opens a screen that lets the user select the query block to add. The Query Blocks are displayed in a tree structure. The top level element (Substance, Mixture, Template) defines the basic document to find, i.e. the user first has to decide if substances, mixtures or templates should be found. When the user selects a query block, a screen is displayed, that allows the user to edit the search criteria. Depending on the selected query block, different screens are displayed.

There are basically four types of queries

- QT1: The queries with a chapter number lower than 4 will find substance, mixture or template documents
- QT2: Queries located under Annotation (2nd level) will find annotated documents
- QT3: The query "Dossier Header" finds dossier header related information
- QT4: The queries with a chapter number from 4 onwards will find endpoint study record and endpoint summary related information. Additionally the Endpoint study records and Endpoint summaries section independent queries will find endpoint study record and endpoint summary related information.



Query Tool screen: The search criteria

The user might leave the search criteria screen empty. If no search criterion is defined, the Query Tool will deliver the following results:

- QT1 queries deliver all substances, mixtures or templates available (depending on the top level element in the query tree).
- QT2 queries deliver all substances, mixtures, templates, endpoint study records or endpoint summaries (depending on the selected query) that have an annotation.
- QT3 queries deliver all dossiers for a substance or a mixture (depending on the top level element).
- QT4 queries deliver all endpoint summaries or endpoint study records for the selected chapter. This can be used to check for an availability of an endpoint for the given chapter.

When search criteria are defined, the result is restricted by the search criteria. The following types of search criteria are available:

- Numerical values (no range): = > < => <=
- Boolean value: is true
- Strings
- Date: = > < => <=
- Pick list/Phrase
- Numerical ranges : Fully including, exact match, overlapping

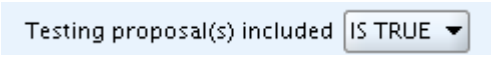
Multiple selections:

For string fields and for pick list fields it is possible to enter multiple values:

- As separator an "|" has to be used.
- The search criteria are limited to a maximum number of 2000 characters.

Boolean values

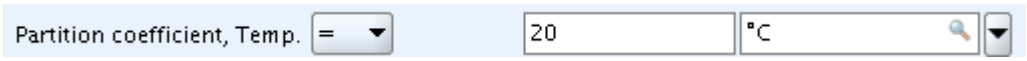
Boolean values can specify only the value "IS TRUE". This searches the index for a field that has the value "true".



Query Tool screen: The boolean search criteria

Numerical values

For numerical values, the value, the search operator and the unit can be specified. The value is automatically converted to the standard unit for the search field (see "5.1 Units"). Note that it is possible for a numeric field to have multiple different standard units. If no unit is specified the unit is ignored, this will only find a value in a standard unit.



Query Tool screen: The numeric value search criteria

Note that there are also numeric values without unit.

String values

String values allow the user to specify multiple values separated with "|". See "5.3 Wildcards and escaping of characters" for details. It is not necessary to specify a search operator for string values. The query tool automatically selects an appropriate search operator based on the input.

Note that the search value is not case-sensitive (except InChl).

Date values

For date values the user can specify a search operator and a search value.

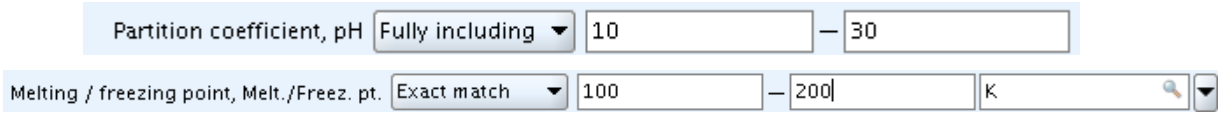
Picklist/Phrase values

Picklist values allow the user to specify multiple values separated with "|". See "5.3 Wildcards and escaping of characters" for details. It is not necessary to specify a search operator for picklist values. The query tool automatically selects an appropriate search operator based on the input. Note that only picklist values in English language are supported and the search value is not case-sensitive.

Numerical range values

For numerical ranges, the range, the search operator and the unit can be specified. The range is automatically converted to the standard unit for the search field (see "5.1 Units"). Note that it is possible for a numeric field to have multiple different standard units. If no unit is specified the unit is ignored, this will only find a value in a standard unit.

Note that there are also numeric ranges without unit.



Query Tool screen: The numeric range search criteria

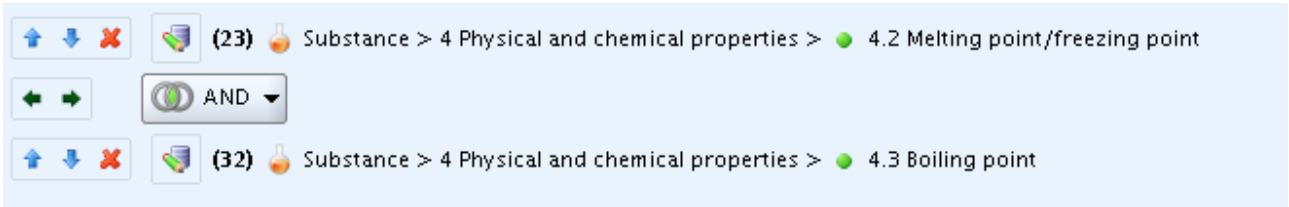
See "5.2 Querying in numeric ranges" for a description of range search operators.

Notes

The user should be aware of the fact that the search operator and the search value have to be specified for a search criterion to be complete. Incomplete search criteria will not be included in the query.

4.2.2 Combining query blocks

The user can add up to 9 query blocks to a query. The blocks can be combined with the operators "AND", "OR" and "NOT".



Query Tool screen: Combining query blocks

The order of the query blocks can be changed by pressing the buttons . The query block can be removed by pressing . Additionally the execution order can be influenced by pressing . The button can be used to edit and change a query block.

The block conjunction logic is based on the unique identifiers (UUID) of the substance, template or mixture in the query result and the unique identifiers of the endpoints of these documents.

Example

	Parent UUID	Child UUID
	e.g. UUID of substance	e.g. UUID of endpoint study record
Block1 (3 hits)	US1	UE1
	US2	UE2
	US3	UE3
Block2 (2 hits)	US3	UE4
	US4	UE5

The AND / OR logic between the blocks will only consider the parent UUID

The result of the conjunction Block1 OR Block2 will be

	Parent UUID	Child UUID
	e.g. UUID of substance	e.g. UUID of endpoint study record
Block1 (3 hits)	US1	UE1
	US2	UE2
	US3	UE3
Block2 (2 hits)	US3	UE4
	US4	UE5

The AND / OR logic between the blocks will only consider the parent UUID

The result of the conjunction Block1 AND Block2 will be

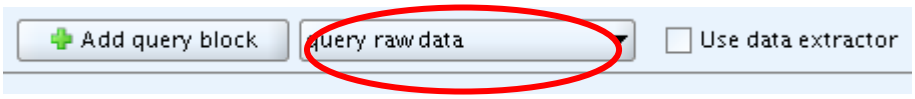
	Parent UUID	Child UUID
	e.g. UUID of substance	e.g. UUID of endpoint study record
	US3	UE3
Block2 (2 hits)	US3	UE4

The result of the conjunction Block1 NOT Block2 will be

	Parent UUID	Child UUID
	e.g. UUID of substance	e.g. UUID of endpoint study record
Block1 (3 hits)	US1	UE1
	US2	UE2

4.2.3 Defining the data layer

The user has to define the data layer in which to search. Possible values are "query raw data", "query dossier data" and "query raw and dossier data".

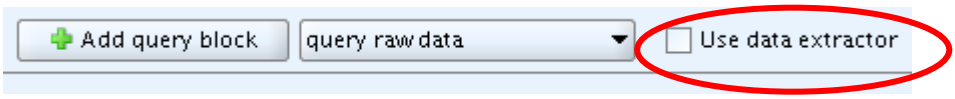


Query Tool screen: Defining the data layer

Depending on the selected value, raw data, dossier data or both is included in the query result. Please be aware that some queries (e.g. "Dossier Header") make only sense when the appropriate data layer is selected, e.g. searching for a dossier header makes only sense when the dossier layer is included in the result.

4.2.4 Defining the query result

The user can define the output of the query result by selecting "Use data extractor".



Query Tool screen: Defining the query result

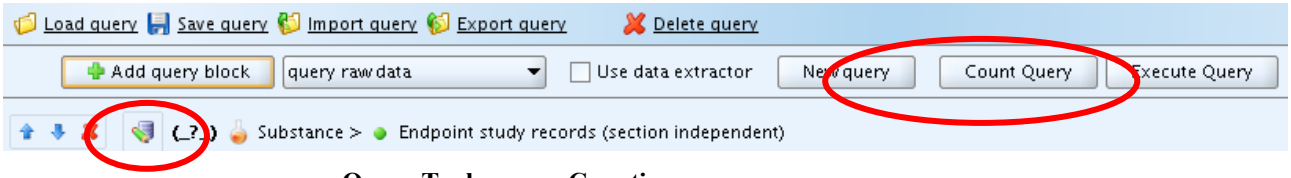
When "Use data extractor" is selected, the query tool summarizes endpoint study records and endpoint summaries in the query result, giving an overview on the complete endpoint. When "Use data extractor" is not selected, the query result shows exactly the fields from the endpoint study records and endpoint summaries for which search criteria were specified. Please note that the data extractor is only available for endpoint study records and endpoint summaries. Substances, mixtures and templates, as well as the dossier header show always the fields specified in the search criteria.

4.3 Count a query

The user should always count a query result before executing a query. The IUCLID 5 Query Tool will only allow counting or executing a query, if the query input is valid. Counting a query result will show an expected result count for each query block. Note that the result count is not always 100% accurate, but always delivers a very good approximation, because it is optimised to deliver a fast response.

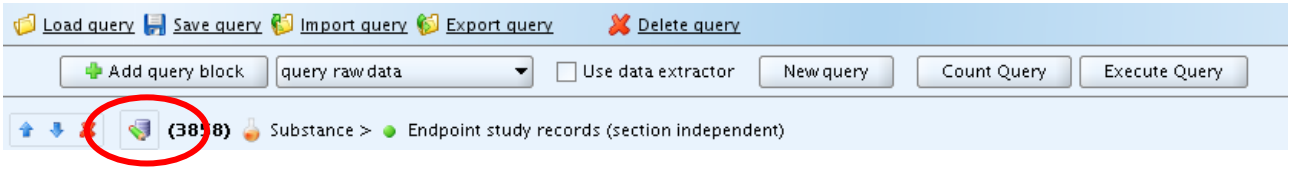
Counting a query result is important, because query results that produce large end results or large block results will take long too run, and will put a heavy load on the servers available memory and processing power. Users should make sure that the query criteria are specified good enough to get only a few hundred hits in the end-result and a maximum of a few thousand hits for the block results. Note that depending on the number of query blocks, the number of concurrent users of the query tool, the available hardware, etc. this recommended numbers will vary.

The screen below shows that a query block displays "(_?_)" before it is counted.



Query Tool screen: Counting a query

When the "Count Query" button is pressed, the block shows the expected result count.



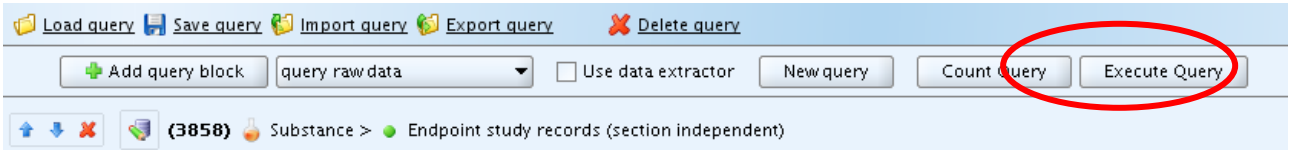
Query Tool screen: Counting a query

Note that the IUCLID 5 Query Tool cannot display a count for the end-result, as this would mean executing the query blocks and combining the query block results. This would take a lot of time, as block combination uses rather complex rules. However, it is a good approach to estimate the final result count by adding up block counts for "OR", using the smaller block count for "AND" and subtracting block counts for "NOT". This allows the user to estimate the size of the end-result.

4.4 Execute a query

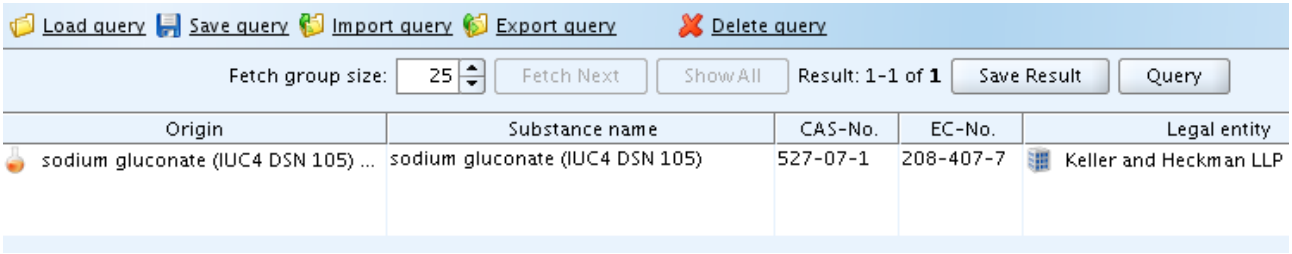
The user should always count a query result before executing a query. The IUCLID 5 Query Tool will only allow counting or executing a query, if the query input is valid.

The user executes a valid query by pressing on the "Execute Query" button.



Query Tool screen: Executing a query

After the query tool has executed the query, the query tool switches to the query result screen.



Query Tool screen: Query results

The user can get additional results – if available - by pressing "Fetch next" or "Show All". The "Fetch group size" indicates the number of items to get when pressing "Fetch next".

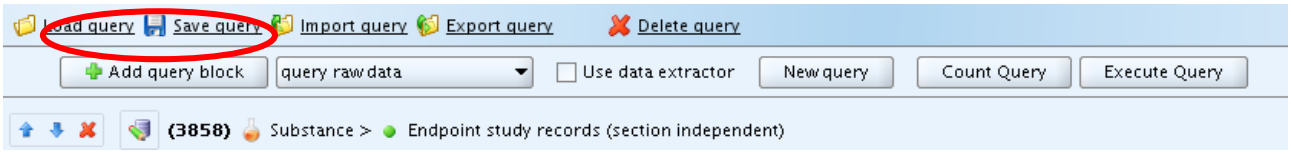
Additionally the user can save the results by pressing "Save Result". The results are saved as CSV format (Character Separated Values), which can be imported into Microsoft Excel or Open Office Calc. The CSV format produced by the query tool uses semicolon as separator and is optimized for Microsoft Excel compatibility.

The button "Query" brings the user back to the query definition screen.

Note that the query can only find IUCLID 5 documents that are included in the index. Depending on the indexer settings it may take from a few minutes up to a day until a search value is included in the index (see "3.2 Configuring the indexer").

4.5 Save a query

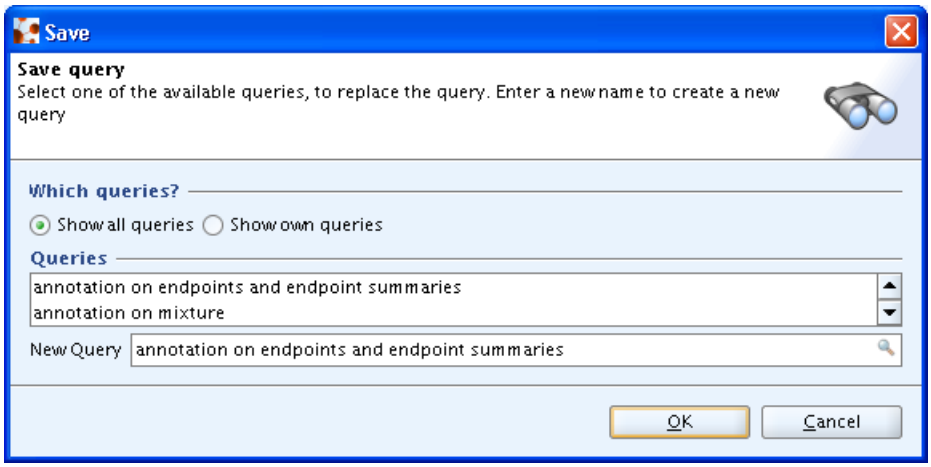
The user can save the queries by pressing the "Save query" button. The query has to be valid to be saved.



Query Tool screen: Saving a query

When the "Save query" button is pressed, a screen appears where the user can enter the name of a query. The user can choose or enter the name of an existing query to replace the query.

The user is stored as owner of a query when the user saves the query. This allows the user to filter the queries by ownership if many queries exist in a system. However, it is possible for every user to take the ownership of every existing query by simply loading and then saving it.

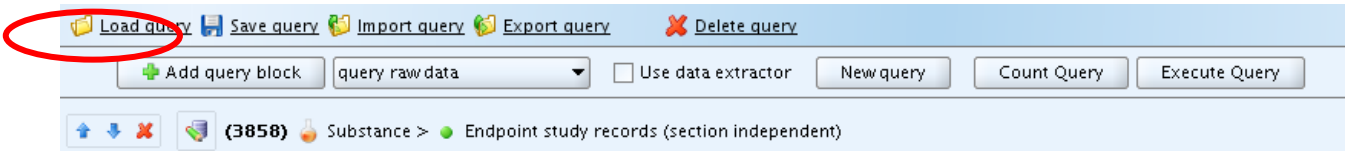


Query Tool screen: Saving a query

Queries are stored in the IUCLID 5 database and are available to every user on the system.

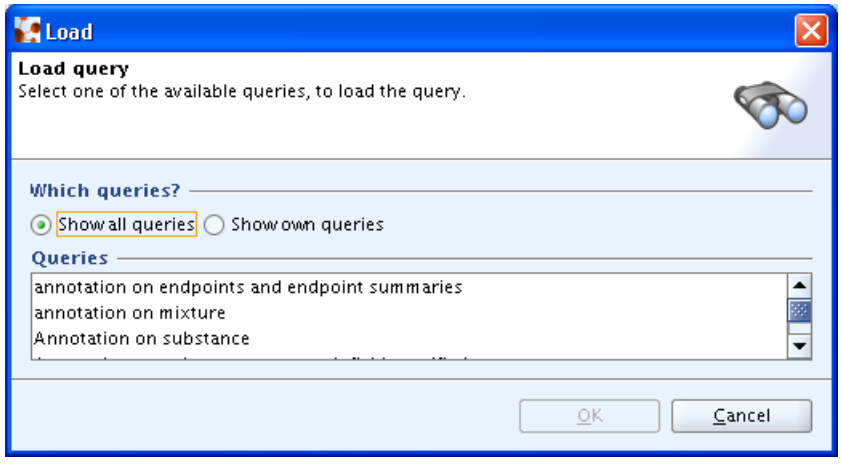
4.6 Load a query

The user can load a query from the database by pressing the "Load query" button.



Query Tool screen: Loading a query

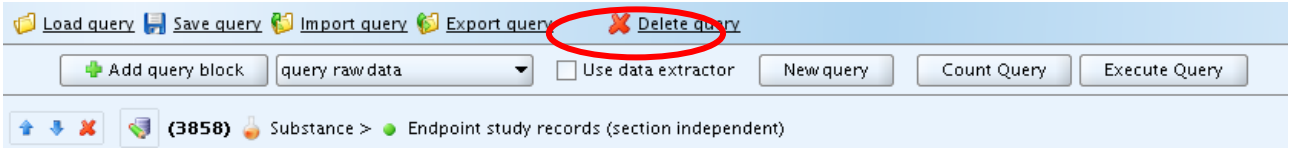
The user can then select the query to load from a list. The list can be filtered by the query owner (show all queries, or only queries owned by the query tool user).



Query Tool screen: Loading a query

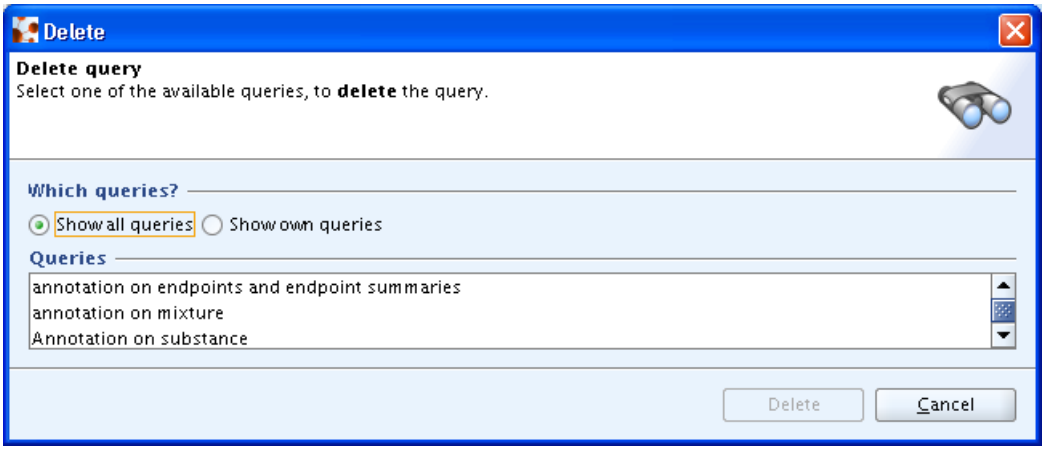
4.7 Delete a query

The user can delete a query from the database by pressing the "Delete query" button.



Query Tool screen: Deleting a query

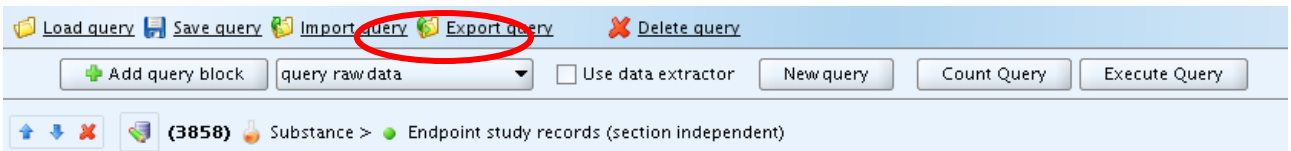
The user can then select the query to delete from a list. The list can be filtered by the query owner (show all queries, or only queries owned by the query tool user).



Query Tool screen: Delete a query

4.8 Export a query

The user can export a query to the file system. Queries are exported to files in XML file format with the extension ".i5q". When the user pushes the "Export query" button, the user can select a directory where to store the export file, and type in a name for the export file. The export operation always exports the query currently shown in the user interface. If the user wants to export a query stored in the database, the user must first load the query from the database and can then export the query.

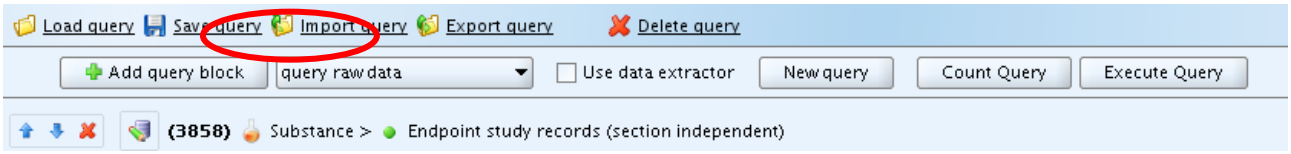


Query Tool screen: Export a query

Users can import the exported query on other IUCLID 5 systems and store them in the database.

4.9 Import a query

The user can import a query by clicking on the "Import query" button. The user can then select the query file from the file system.



Query Tool screen: Import a query

Imported queries are not stored in the IUCLID 5 database until the "Save query" button is pressed. This allows the user to build a private collection of queries without making the queries available to other user. Additionally, users are able to edit imported queries before publishing them (by storing the queries in the database).

4.10 Important notes

The user builds queries that search in the index created by the query indexer. Users should be aware, that the index is updated in a configurable time interval. So, changes made to IUCLID 5 documents will not be instantly reflected. Before a field can be found, the indexer has to run and this field has to be stored in the index.

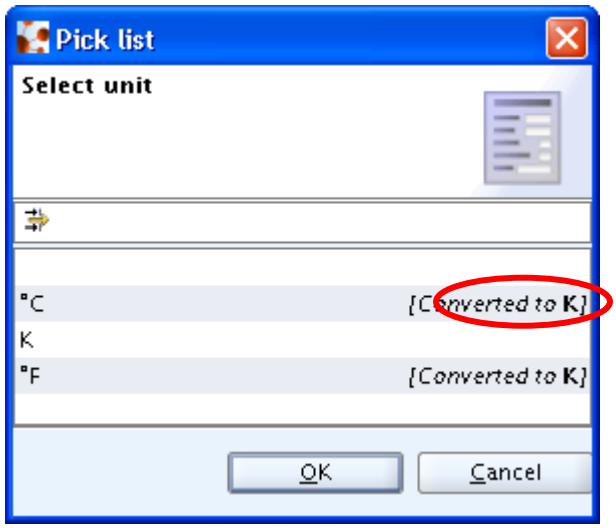
The IUCLID 5 Query Tool does not restrict the result sets that might be returned from the query in any way. However, the result set is restricted by the available resources on the IUCLID 5 client and the IUCLID 5 server (or the IUCLID 5 workstation for the standalone installation). When the result set is excessively large (thousands of result items), the system might encounter an out-of-memory situation which will happen at a random position (not necessarily inside the Query Tool).

5 Appendix

5.1 Units

The indexer of the IUCLID 5 Query Tool converts all numeric values that have a unit to a defined standard unit. When the user enters a query value and specifies a unit, the query value is also automatically converted to the standard unit.

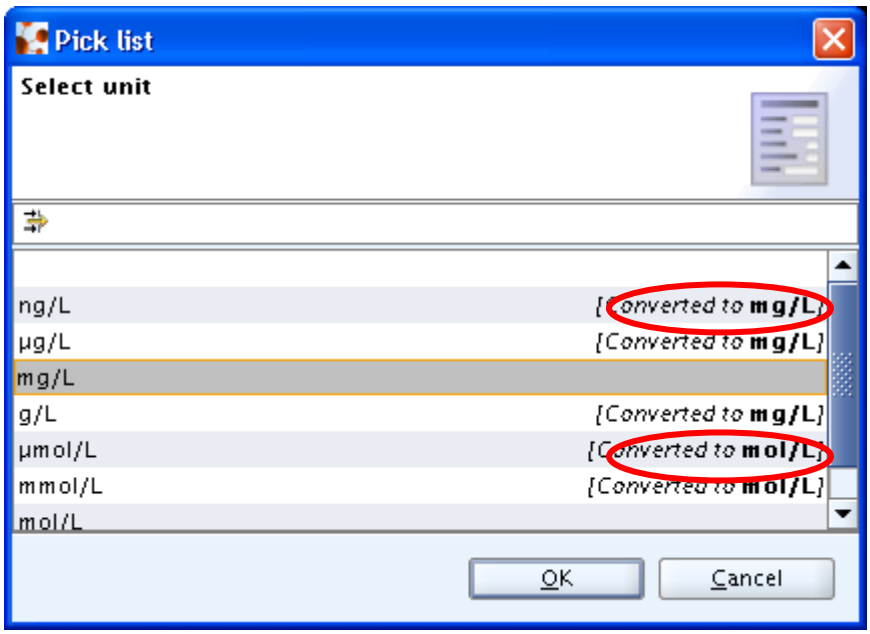
When the unit is selected for the query criteria, the query tool screen shows the standard unit to which values are converted.



Query Tool screen: Standard units

The standard units are displayed to the right of the selected unit.

Note that there are lists of units where more than one standard unit is contained.



Additionally there might be "other:" units in the unit list. If "other:" is selected, the user has to specify the unit exactly as contained in the searched document. As the indexer does not understand the meaning of "other:" units, no conversion to standard units is done.

5.2 Querying in numeric ranges

In IUCLID numeric values are often entered as ranges. When querying for numeric values the user has to specify an upper and a lower values. The user's specification controls how two numerical intervals are compared with each other.

5.2.1 Ranges in IUCLID 5

A range in IUCLID 5 lets the user specify a lower qualifier, a lower value, an upper qualifier, an upper value and a unit. The interpretation of the range depends on the filled fields.

Numeric range fields typically comprise the following subfields:

- Qualifier (lower value): The lower qualifier field provides a list with following qualifiers: >, >=, and ca. No "equal to" sign (=) is provided. Instead, this is indicated by no operator being selected, i.e. the qualifier field is simply left empty.
- Numeric field (lower value): The lower value of the range. The lower value is also used for entering a single numeric value.
- Qualifier (upper value): The upper qualifier field provides a list with following qualifiers: <, <=, and ca. No "equal to" sign (=) is provided. Instead, this is indicated by no operator being selected, i.e. the qualifier field is simply left empty.
- Numeric field (upper value): The upper value of the range. The upper value is also used for entering a single numeric value.
- Unit field: The Unit field is a list field for entering the unit of measurement. In the case of an open list field, an associated text field follows next to it, where you can specify any "other:" unit. Please be aware of the fact that specifying a unit that is not included in the unit list will prevent the IUCLID 5 query tool from converting the range to a standard unit, and make searching for these ranges nearly impossible. Some numeric ranges come without a unit field either if the unit is indicated in the field label or if the value is dimensionless.

Numeric range fields provide enough flexibility to enter either a single numeric value (and the unit) only, or a single value preceded by a qualifier or two numeric values, preceded by qualifiers or none. The different lists provided in the lower and upper qualifier fields and the underlying control rules ensure that no invalid numeric ranges can be entered. Following conventions apply:

- In the lower qualifier and numeric field, enter a numeric value preceded either by no qualifier, ">", ">=" or "ca." (e.g. "20", ">20", ">=20", "20")
- In the upper qualifier and numeric field, enter a numeric value only
  - if either a lower value is already entered to specify a numeric range

  or

  - if the upper value is preceded by either operator "<" or "<=".

Example	Range (graphical representation)
154 – 170	154 ◄——► 170
> 154 – 170	154 ——► 170
> 154 – < 170	154 ——— 170
> 154	154 ———► ∞
< 170	-∞ ———► 170
154	154 ►
ca. 154	154 ►

5.2.2 Specifying the search criteria

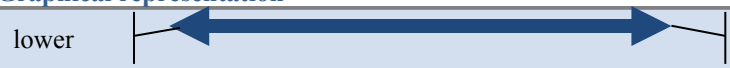
The range search criteria are specified by providing a lower value, an upper value, a unit and a search operator.

Query Tool screen: Searching ranges

The criteria entry fields let the user specify the search criteria for ranges, with the least possible input. The user does not have to specify the qualifiers of the IUCLID 5 ranges. Searching for "lower value – upper value" always finds ranges "lower value – upper value", ">= lower value – upper value", "> lower value – upper value", etc.

Depending on the entered values, the search is constructed.

- Only lower value specified: Search range is lower value up to infinity
- Only upper value specified: Search range is upper value down to negative infinity
- Both values specified: Search range is lower value to upper value
- To specify one single value, enter the same value in the lower and upper search field.

Search criteria	Graphical representation
lower – upper	
lower	
upper	
Lower = upper	

Together with the search operators ("overlapping", "fully including" and "only exact match"), this is sufficient to specify exactly which IUCLID 5 range to find.

Examples:

IUCLID 5 range	Range criteria	Finds range
154 – 170 or > 154 – 170 or > 154 – < 170 or > 154 or < 170 or 154	Overlapping 154 – 170	yes
154 – 170 or > 154 – 170 or > 154 – < 170 or > 154 or < 170	Overlapping 154 – 155	yes
154 – 170	Fully including 154 – 155	no
154 – 170	Fully including 152 – 170	yes
154	Fully including 154	yes
154	Overlapping 154	yes
154	Exact match 154	**no** (The search criteria tries to find a range from 154 to infinity)
154	Exact match 154 – 154	**yes** (The search criteria tries to find a range from 154 to 154)
< 170	Exact match 170 – 170	no
< 170	Exact match – 170	yes
< 170	Fully including 170 –	no
< 170	Fully including – 175	yes

5.2.3 The range search operators

Range searches offer the operators "overlapping", "fully including" and "only exact match".

Sample: Ranges overlapping

	entered data	*graphical illustration*
A search for	**10 .. 20**	0 ----------30
finds	20	0 ----------30
	5 .. 15	0 ----------30
	15 .. 20	0----------30
	> 18	0 ----------30
	< 12	0 ----------30
	7 .. 25	0 ----------30

Sample: Ranges fully including

	entered data	*graphical illustration*
A search for	**10 .. 20**	0 ---------- 30
finds	20	0 ---------- 30
	15 .. 20	0 ---------- 30
but does not find 5 .. 15		0 ---------- 30
	> 18	0 ---------- 30
	< 12	0 ---------- 30
	7 .. 25	0 ---------- 30

Sample: Ranges only exact match

	entered data	*graphical illustration*
A search for	**10 .. 20**	0 ---------- 30
finds	10 .. 20	0 ---------- 30
finds	>10 ..<20	0 ---------- 30
but does not find 20		0 ---------- 30
	15 .. 20	0 ---------- 30
5 .. 15		0 ---------- 30
	> 18	0 ---------- 30
	< 12	0 ---------- 30
	7 .. 25	0 ---------- 30

5.3 Wildcards and escaping of characters

The IUCLID 5 Query Tool supports three different wildcard on querying string and picklist fields:

Wildcard	Usage	Sample
*	No, one or multiple characters	**Meth* ➔ Methan, Methen, Methanol**

	(windows like)	
%	No, one or multiple characters (SQL like)	**Meth% ➔ Methan, Methen, Methanol, Meth**
_	Exactly one character (SQL like)	**Meth_n ➔ Methan, Methen**

Beside of the three above mentioned wildcards, the separator ("|"), used to enter multiple values has to be escaped, too.

If the user would like to search for a string, containing one of the four special characters, he has to enter the character twice:

enter "%%" to search for "%"

enter "__" to search for "_"

enter "**" to search for "*"

enter "||" to search for "|"

European Chemicals Agency

IUCLID 5 Support

Query Tool plug-in Installation and User Guide
January 2009 v 1.0

http://iuclid.eu